DOCUMENT RESUME

ED 131 074 SP 010 635

AUTHOR Koran, John J., Jr.; Koran, Mary Lou
TITLE Validating a Teacher Behavior by Student Performance.
INSTITUTION Florida Univ., Gainesville. Inst. for Development of Human Resources.
SPONS AGENCY Florida State Dept. of Education, Tallahassee.
REPORT NO FSDE-730-063
PUB DATE 73
CONTRACT 730-063
NOTE 34p.; Research paper submitted to Association of Teacher Educators, 1975

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Academic Achievement; Comparative Analysis; *Comparative Testing; *Educational Research; Learning Theories; Preservice Education; *Questioning Techniques; Skill Development; *Teacher Education; Teaching Methods; Teaching Techniques; *Test Results

ABSTRACT

In an investigation of the relationship between teacher analytic questioning skill and student performance, 69 preservice teachers were assigned to one of three treatment conditions designed for questioning training: (1) written model, protocol form; (2) written model, transcript form; (3) placebo, control group. Two hundred ninety-five eigth grade students were assigned to groups of three to four to act as students in a teaching setting. Both students and teachers were given a communication to read as the basis for a twenty-minute lesson. After the lesson, teachers were tested on their ability to identify analytic questions, and students were tested on lesson content, identification of analytic components, and one week later their ability to analyze a different communication. A one-way analysis of variance was used to test the effects of the treatments on the acquisition of analytic questioning skill by the teacher trainees and also on the analytic responses of the students. Performance of subjects in the treatment groups significantly exceeded those in the control group on: (1) the teacher and student identification of analytic questions; (2) teacher frequency, quality and variety of analytic questions; and (3) student frequency and variety of analytic responses. Multiple linear regression analysis showed a significant relationship between frequency, variety and quality of teacher analytic questioning behavior, the frequency and variety of student analytic responses and also student performance on the written identification of analytic questions. The best combinations of teacher behavior for predicting student performance varied for different student learning outcomes. (Author/MB)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED131074

12

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Validating a Teacher Behavior by Student Performance[1]

John J. Koran, Jr.
Professor of Education
Division of Curriculum and Instruction

Mary Lou Koran
Associate Professor of Education
Department of Educational Foundations

Institute for the Development of Human Resources
College of Education
University of Florida
Gainesville, Florida 32611

SP010635

[1]A research study funded by the Florida State Department of Education, Research and Development Division, Contract Number 730-063, December, 1973.

Validating a Teacher Behavior by Student Performance

Introduction

Traditionally teacher training in the methodology of teaching has implicitly or explicitly included the acquisition of particular skills thought to be related with student performance (Combs, A. W., 1972; McDonald, F. J., and Allen, D. W., 1967; McDonald, F. J. and Koran, M. L., 1969; Koran, J. J. Jr., Koran, M. L., and McDonald, F. J., 1972). However, with the movement towards accountability and performance based programs for teachers, (Elam, S., 1971; Elam, S., 1972; Houston and Housam, 1972) educators and the private and political sector are calling for concrete evidence of the relationship between what a teacher is taught to do and say in a training program, and the ultimate effect on students working with that teacher. Rosenshine (1970) dramatizes both the paucity of research of this type, and the need for it, in his review of literature on the relationship between teacher behavior and student achievement. He emphasizes that most of the researchpperformed and reviewed in this area has been poorly designed or executed and inconclusive with respect to student gains (Rosenshine and Furst, 1971; Rosenshine and Furst, 1973). This has prompted Rosenshine and Furst (1971) to propose a five phase program to generate useful data in studies of this type: Essential components or instructional variables considered specific to a curriculum should be identified. Teachers should then be trained to use these instructional variables properly. The relationship between instructional activities and behavioral changes of students should be identified. Modification of training procedures and/or materials should be made on the basis of the latter phase. Finally, new research with appropriate controls on training procedures and/or materials

should be undertaken. Rosner and Kay (1974) have abbreviated this sequence in their attempt to schematize the structure for arriving at competency based teacher education. (Figure 1)

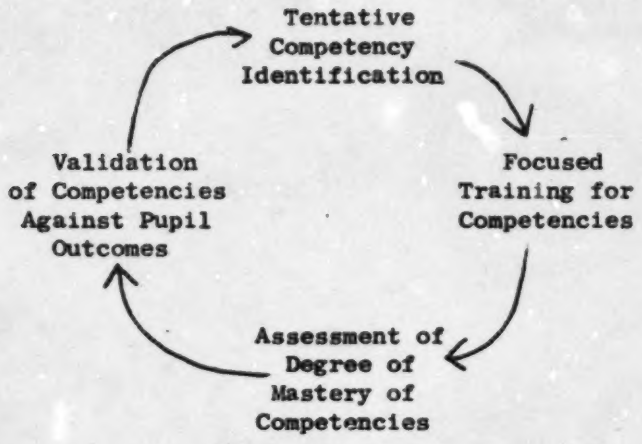
Tentative Competency Identification
Focused Training for Competencies
Assessment of Degree of Mastery of Competencies
Validation of Competencies Against Pupil Outcomes

FIGURE 1

While both of these approaches share common components, one element they neglect by emphasizing teacher training to criterion and then validation of competencies against pupil outcomes is the inability of this model to study the effects on students of teachers who have not been trained to criterion. This paper will describe an alternative two phase model for doing both training research and subsequent validation in terms of student performance which incorporates training teachers to a range of skill levels and examining the effects on school students of teachers who have mastered the skill to both high and low levels of performance.

Theory and Design

This study can be divided into two phases. In the first phase training methods are compared to assess their effects on the acquisition of teacher skill in analytic questioning. Any training procedures that have theoretical

and/or empirical support could be used in this phase to compare their effects as training methods, and to produce a range of teacher proficiency on a given skill. In this phase the training method is the independent variable and the teacher behavior the dependent variable.

In the second phase of this study the teacher behavior acquired as a result of the treatments is related to student learning. Here, teacher behavior becomes the independent variable and student learning the dependent variable. This two-phase design has the potential for yielding maximum data when testing the effects of different training methods on teachers as well as the effects of subsequent teacher behavior on student learning. (Table 1.)

Insert Table 1 approximately here

The procedure was selected because the apriori setting of criterion levels for teacher acquisition of a skill or student learning seems to make little sense when so little is known about most skills, and arbitrarily set criteria may mask practically important relationships between teacher and student behavior. When different treatments and a control are used there is a high probability of obtaining a range of teacher behavior from high to low performance. If the desired teacher behavior is, in fact, related to student learning, then learning may be expected to vary correspondingly, and the combined data could lead to the setting of empirically validated criterion levels based on this information.

Data of this type can also lead to information regarding whether the independent and dependent variables are related in a linear or non-linear

way, again providing evidence to suggest potential criterion levels for similar samples with this skill. Moreover, the setting of teacher criterion performance in terms of its effect on student learning must necessarily consider the multivariate nature of learning outcomes. The fact that learning may be demonstrated in a number of different ways which are far from perfectly correlated has been well documented. Any specific treatment variable or teacher behavior may have multiple effects, and that which is best in producing immediate mastery is not necessarily best for delayed performance, transfer, affective outcomes or other indices of achievement. Ultimately the establishment of criteria for teacher performance in terms of student learning must necessarily consider which specific student learning outcomes are to be obtained. This must be empirically determined by exploring the effects of a range of different teacher behaviors on multiple student learning outcomes, rather than established on an apriori basis.

In this study, the theoretical basis for the selection of treatments was social learning theory (modeling, or observational learning) Bandura and Walters (1963), Bandura (1969), Bandura (1973). Both theory and accompanying research support the contention that simple and complex behaviors may be acquired or modified through observation with no direct external reinforcement. Modeling has been found to be more effective than operant conditioning in transmitting new response patterns (Bandura and McDonald, 1963), with the provision of a model alone being as effective as the combination of modeling and reinforcement for initial learning. Other research has shown that film-mediated models have been as effective as live models in producing the behavior change (Bandura, Ross, and Ross, 1963).

The implications of these findings for teacher training is that the provision of live or symbolic models (written models) displaying desired

teacher behaviors may provide an effective alternative to traditional verbal-descriptive techniques of training (McDonald, F. J. and Allen, D., 1967). Applications of modeling procedures to teacher training have demonstrated the efficacy of both film-mediated models (Koran, J. J. Jr., 1969; Koran, M. L., McDonald, F. J., Snow, R. E., 1971; Koran, J. J. Jr., Koran, M. L., McDonald, F. J., 1972) and written models of the type used in this study (Koran, J. J. Jr., 1970; Koran, J. J. Jr., 1971) for the development of questioning skills in preservice teachers. The general superiority of film-mediated models over written models, and the intersection of individual differences with various modeling procedures in the acquisition of teaching skills (Koran, M. L., Snow, R. E., McDonald, F. J., 1971) has been shown on at least one occasion. Other research (Masters and Branch, 1969) supports the efficacy of exploring a variety of written models for inducing different types of teacher behavior change. Since variations on a symbolic (written) model have been shown to be effective methods of influencing some teacher behaviors and they are inexpensive, portable, and permit greater control over environmental conditions in the school and university setting than live teacher models or videotape models, this type of training method seems to warrant closer scrutiny. For these reasons and the fact that materials of this type fit the requirements of this study they were selected as training methods.

It was anticipated that Ss in the two treatment groups would produce significantly more analytic responses than those in the control group on both the written and microteaching measures and that all three groups would produce the necessary range of teacher behaviors to permit exploration of the effects of teacher acquisition of a skill to high and low levels of

proficiency, and student learning. Accordingly, it was expected that teachers who achieved the skill of asking analytic questions to a high level of proficiency would influence their students to more analytic thinking and consequently greater analytic performance than teachers at the other end of the continuum.

Methods

Subjects

The experimental sample consisted of 69 preservice secondary teachers nearing the end of their professional sequence of education courses and approximately 295 eighth grade students from a large middle school. Since the learning task during the microteaching lesson was the analysis of a written communication, students who were judged to be non-readers on the basis of reading achievement scores and teacher recommendation were excluded from the pool of students from which microstudents were randomly selected for microteaching groups.

Materials and Procedures

Two sets of training materials were used: a written model: protocol form which can be distinguished from other forms of written instruction in that it operationally defined analytic question categories (Bloom, 1956) and gave examples of teacher analytic questions for each category: a written model: transcript form which included a written transcript of a dialogue between a teacher and four students analyzing the communication "When We Are Gods" by Archibald McLeish. Inserted in this transcript preceeding each analytic question was a written cue, in parentheses, advising the reader to note "how the teacher asked" a certain type of analytic question.

A third group of Ss received initial set induction materials describing the general purpose of the microteaching lesson and the article, "Was Thoreau a Hippy," which was the basis of the microlesson, and was told to teach a lesson analyzing the article. This same article was the basis of the lesson for the other two groups.

Insert Table 2 approximately here

Table 2 describes the general design and treatment procedures. All microteaching sessions were 20 minutes in length and took place in two rooms in a school. The 295 microstudents were randomly assigned in groups of four to each microteaching session. Treatments were randomly assigned to subjects as they arrived at the school. The entire study took place over a period of seven days. Each treatment was randomly distributed over days and times. After the audio-recorded microteaching lesson was completed both the teachers and students took written tests to measure recognition of analytic questions. The microstudents took additional tests to assess their acquisition of the content of the lesson, their affective reaction to the lesson, and one week later a transfer test requiring analysis of a similar communication.

The recorded teacher-student interactions were rated by three raters for the frequency, variety, and quality of analytic questions asked. Frequency counts included the total number of analytic questions asked by the teacher or responses given by the student. Variety was defined as the total number of different analytic questioning categories the trainee used and the students responded in. Quality, high or low was used to distinguish

between near versus more remote approximations to the behaviors illustrated in the symbolic model. Information content was selected as the dimension along which approximation to the behavior would be assessed. In both cases, the questioning behavior exemplified in the written materials required the student to supply the maximum amount of information that could be elicited by that particular type of question. The closest approximations to the illustrated behaviors would also do this. Accordingly, high quality questions were those which elicited from students all information relevant to answering them. Questions requiring students to supply only part of the information relevant to answering them, or requiring only that students agree, disagree or select from among given alternatives would be low quality questions. The type of criterion measure and the reliability of each measure are reported in Table 3. Reliability of the written measure was determined

---

Insert Table 3 approximately here

---

using the Cronbach Alpha. The reliability of the scores derived from the twenty-minute tape was determined by three trained raters after approximately thirty hours of training and ninety hours of rating. Reliability coefficients were derived for the frequency, variety and quality of the criterion behaviors.

An analysis of variance repeated measures model described by Winer (1962) was used to analyze the reliability of rater 1, 2, 3 on tapes X_1-X_n. This analysis provided the reliability of the mean rating of the three raters. The reliability of the mean of the three ratings ranged from .89 to .99. The estimate of the reliability of a single rating for adjusted data had similarly high reliability. Since these coefficients are extremely high they suggest that the criterion measures are sufficiently reliable for

testing treatment effects in phase one of this study and teacher-student performance relationships in phase two.

In summary, audio tape data provided information regarding the level of acquisition of the teacher behavior, asking analytic questions, and the degree of student responsiveness. The former data is a test of the training procedures effectiveness on teachers. The latter data is an indication of the effects of the teacher behavior on students. Written measures on both teachers and students produced other data which could be related in the same way to contribute to making inferences about the effects of acquired teacher behavior on a variety of student performance.

Results

It will be recalled that there are two phases to this study. The first was to assess the effects of the different training procedures on the acquisition of teacher analytic questioning skills as a result of exposure to two symbolic modeling treatments, and to assess the relative effects of these treatments on student analytic responses. The second phase was to further explore the relationship of teacher performance to student performance. Two general hypotheses were proposed: (1) that the two treatment groups would produce significantly greater teacher behavior change than the control group; and (2) that the microstudents would produce a significantly greater frequency of analytic responses than those in the control group on both the written and the microteaching measures.

Treatment Main Effects

The first hypothesis was tested using a one-way analysis of variance for the frequency, variety, and quality of acquired teacher behavior across

treatment groups on the microteaching and also on written performance measures. Similar analyses were used for the student microteaching and

---

Insert Tables 4, 5, and 6 approximately here

---

written performance measures. Table 4 reports the means and standard deviations, by treatment, for both written and microteaching measures. Tables 5 and 6 report the analysis of variance results. These tables show a significant difference ($p < .01$) between groups on the teacher written criterion measure. Comparison between pairs of treatments showed that both modeling treatment groups performed significantly better on the written measure than the control group ($p < .01$). Moreover, the written model: transcript form was significantly more effective ($p < .05$) than the written model protocol form in producing behavior change. On the parallel written measure for students the analysis of variance indicated similar significant differences ($p < .01$) with the treatment groups exceeding the control group on the performance measure. However, there were no differences in student performance on the written measure between the two modeling treatment groups. There were also no significant differences between the treatments and the control on the student content measure and the student written transfer measure (Tables 6 and 7).

Insert Tables 7 and 8 approximately here

Table 8 summarizes the results of a series of one-way analysis of variance tests on both teacher and student analytic behavior. In all cases, the treatment groups produced significantly greater acquisition of the criterion behavior in teachers than the control group ($p < .01$), and correspondingly, significantly more analytic responses of school students

who were in microteaching groups with teachers who acquired to higher levels, the skill of asking analytic questions ($p < .01$). The protocol materials generally produced greater acquisition of the criterion behaviors than the transcript model, with significant differences shown in four of six categories: total number of questions asked by the teacher ($p < .05$); number of low quality questions asked by the teacher ($p < .01$); total number of student analytic responses ($p < .01$); and number of categories of student responses ($p < .05$).

Insert Table 9 approximately here

Table 9 summarizes the student responses to a six item instrument designed to assess some affective dimensions of the student experience. These data are reported in frequency form for each treatment and item. Some generalizations which might be derived from the student responses here are:

1. Students participating in the study reported that the lesson on the communication "Was Thoreau a Hippy" was either very interesting or of some interest.
2. Students reported the skill of analyzing a communication in terms of its component parts was important and/or required in order to understand it.
3. Students reported the questions the trainees asked were either of some help or very helpful.
4. Students reported that the questions the trainees asked either required some thought or forced considerable thought.
5. Students reported that the lesson was satisfying "as it was" rather than with more questions or fewer questions asked.

6. Students reported that in comparing this lesson with their "usual lessons" they felt they either learned as efficiently or more efficiently.

From the foregoing data it could be concluded that in phase one of this study the treatments did indeed have a significant influence on teacher acquisition of recognizing and asking analytic questions as measured by the instruments and rating procedures used. Students exposed to teachers who received the treatment conditions could recognize analytic questions and discriminate between analytic and non-analytic questions significantly better than students exposed to the control teachers. They could also make significantly more analytic responses in significantly more categories than the control students. From this data alone one can infer the efficacy of the treatments on teacher acquisition of this behavior and the high correlation between teacher behavior and student behavior. Thus, both research hypotheses were supported.

Finally, microstudents who participated in the study tended to respond positively to six items on a questionnaire to assess their interest and receptiveness to questioning strategies (Table 9).

Teacher-Student Relationships

In order to further explore the nature of the relationships between the teacher and student variables, correlation and multiple regression procedures were used. The correlation of independent and dependent

---

Insert Table 10 approximately here

---

variables is shown in Table 10. The results for the four microteaching

performance measures previously described were analyzed separately although it can be seen that in some instances the correlations among these measures were sufficiently high to suggest that they may not represent psychologically different variables. This was done because of the somewhat different relationships to student behavior observed in some instances, and because of the general unavailability of computer programs for multivariate analysis of variance techniques. However, for practical purposes such results may indicate that the number of variables to be considered in establishing teacher criterion performance may conceivably be reduced.

While there are a large number of statistically significant relationships between teacher performance and student behavior, it should be recognized that due to the size of the sample a particular correlation may be statistically significant, while practically speaking it accounts for very little performance variance.

Multiple regression analysis using stepwise regression procedures was employed to determine the best combinations of teacher behavior in predicting a number of different student learning outcomes. These results are shown

---

Insert Table 11 approximately here

---

in Table 11. They indicate that the total frequency of analytic questions, variety, low quality of analytic questions, and teacher performance on the written measures strongly predict ($p < .01$) frequency of student analytic responses, accounting for 96% of the performance variance. Similarly the frequency and variety of analytic questioning behaviors are highly significant predictors ($p < .01$) for the variety of student analytic response categories, accounting for 90% of the performance variance. In addition, the variety of categories of analytic questioning used by the teacher was a

significant predictor ($p < .01$) of student ability to identify categories and types of analytic questions on the student written measure, although only 4% of the total variance was accounted for. There were no significant predictors for either the student content or transfer tests.

Since the amount of variance accounted for in the frequency and variety of analytic student responses was so high, efforts to ascertain the sources of the additional 4 to 10 percent variance will not be reported here. Moreover, additional efforts did not produce a significant increase in prediction for student identification of analytic questions in the written measure, transfer or content tests.

Thus, it can be seen that some measures of student performance were strongly related to various aspects of teacher analytic questioning skill, while others were not. The best combinations of teacher behavior for predicting student performance varied from variable to variable, with the variety of categories of analytic questioning used by the teacher serving as the most universal predictor followed by the total frequency of analytic questions used by the teacher.

Discussion

The major purpose of this experimental study was to train teachers to various levels of performance on a teaching skill and to assess their differential effects on student learning. Two written models were tested against each other and a control group to ascertain their potential as teacher training methods. In each model the amount and type of information communicated varied, and patterns of information processing were required that may be more or less demanding and of more or less value as training strategies. The protocol model incorporated both general and specific examples of the teacher behavior to be acquired. Its general superiority

over the written transcript in this study, although not unexpected, is somewhat more dramatic than expected. Although research evidence is available (Masters and Branch, 1969; Koran, J. J. Jr., 1970, 1971; and Bandura, 1973) supporting the effects of the protocol type model, a previous study on the acquisition of a teaching skill (Koran, J. J. Jr., Koran, M. L., McDonald, F. J., 1972) has reported the positive effects of having student responses present in a modeling treatment. While the transcript model presented the student-teacher dialogue, the presence of student responses apparently was not strong enough to overcome the specificity of information contained in the protocol model, even though student response elements were completely absent in the protocol. In addition, the protocol provided a wider range of types of analytic questions to be inferred by the learner. This characteristic could explain significant differences between the treatments in frequency, variety and quality of questions generated by Ss during microteaching. Variety and number of categories of analytic questions would appear to be those areas where significant differences ($p < .05$) would appear between the protocol and the transcript model. More analytic questions were asked in the protocol model treatment than the transcript treatment suggesting that the operational definitions and specific examples provided more specific information, while the transcript model provided more limited information from which inferences about the behavior could be made. This same protocol model specificity of the criterion behavior could explain the significantly higher number of low quality questions asked by the teacher ($p < .01$). Since the transcript model provided generally high quality prototype questions, inexperienced trainees probably found it difficult to go beyond the example to generate

lower level examples. Correspondingly, the total number of student analytic responses were significantly affected ($p < .05$) and number of categories of student responses was significantly higher for the protocol over the transcript ($p < .05$). This supports the contention that students exposed to teachers who employ the criterion behaviors also acquire them to a level that is significantly greater than students taught by teachers who do not use these behaviors.

Moreover, Fuller (1959) has posed a developmental conceptualization of teacher concerns in which early teaching concerns are believed to focus on content adequacy, class control and superior evaluation, while only toward the end of student teaching do concerns focus on student learning and self evaluation. Since the particular trainees in the sample had had no previous microteaching or student teaching experience, it may well be that the protocol model was best suited to this inexperienced sample in providing a wider range of specific examples which they could use in their lesson, while more experienced teacher trainees would profit more from the model in which the effects of teacher behavior on student behavior could be observed. This latter exploration could account for the superiority of the transcript form over the protocol form on the one written teacher measure. These alternative explanations should certainly be explored in future research. However, for the purposes of this study this teacher sample did acquire the criterion behaviors to a significantly higher level from the treatment groups than the control group.

That students who were in microteaching groups with teachers who had a treatment also showed significantly greater acquisition of the criterion behavior than the control students is not surprosing. If a teacher asked

an analytic question of his microteaching group the students had to engage in analytic thinking in order to answer. If they could not answer, the teacher probably explained what the question category was and/or rephrased the question so as to clarify its intent. Teachers in the treatment groups had a wider range of analytic questioning behavior than control teachers. In addition, with the asking of the question the teacher modeled the type of questions one should ask oneself during the analysis of a communication. Student correct responses were likely to be reinforced during the lesson while incorrect responses received corrective feedback and additional practice. Under these conditions it is not unlikely that students would have a significantly higher frequency and variety of analytic responses when they were taught by a trained teacher who had acquired the skill to a higher level of proficiency rather than by a control teacher or a teacher who had limited command of the skill.

The fact that the written test requiring students to identify and discriminate analytic from non-analytic questions was significantly superior for the treatment groups as compared with the control probably is related to previous explanations. Students who were not exposed to teachers who exhibited the criterion behavior had little, if any, experience with analytic thinking and making analytic responses, thus little practice and little feedback on the criterion behavior. Hence students in the treatment groups exceeded those in the control group. Or to put it another way, when the teacher was trained on the skill to various levels of performance, his/her students reflected this in their own recognition of the categories of the behavior on a written test.

It is interesting to note that content acquisition on the communication and transfer of the analysis of a communication skill after a week did not significantly differ between treatment and control groups. Since all students had an opportunity to read the communication, as the teacher did, prior to microteaching it is likely that knowledge was acquired during this experience. In order to test this, a sample of forty-two students which simply read the material without participating in microteaching was randomly selected and tested on the content of the treatment communication. They did not differ significantly from the other groups on content acquisition, but did differ significantly ($p < .01$) on recognition of the components of analytic thinking from the treatment groups. It would appear from the foregoing data that analyzing a communication according to its parts is substantially different from acquiring the content of the communication and does not necessarily guarantee the latter. This tends to be supported by the low correlation between the two measures in Table 10.

With regards to transfer, the analysis of elements skill is probably sufficiently complex that a twenty-minute lesson could well have been too short to produce meaningful transfer. At the same time, the behavior was sufficiently novel that microstudents could have shared considerable information during the time lapse between treatment and transfer test to result in no difference between treatments on the transfer measure. If indeed analytic thinking has outcomes other than content acquisition and the retention measure was content centered, one would expect data consistent with the posttest data. Unfortunately no measure was made on a retention test of the ability to recognize analytic components. If this study were replicated, or in future studies of a similar nature, it would appear wise to expose

students to analytic questioning over a longer period of time and to administer a transfer test covering both analytic thinking and content.

Correlational and multiple regression analyses (Tables 10,11) tend to support and illuminate the relationships shown in the analysis of variance results. These data support the contention that a number of student performance outcomes were highly correlated with the analytic questioning behavior of their teachers. Frequency, variety and quality of teacher analytic questioning behavior was strongly related to the frequency and variety of student analytic responses, and the variety of teacher analytic questioning behavior was also significantly related to student ability to identify categories and types of analytic questions. However, teacher analytic questioning behavior was not related either to student content or transfer tests, quite possibly for the reasons previously discussed. Multiple regression analysis further supported the expectation that different combinations of teacher behaviors were significant predictors for different student learning outcomes. Therefore, teacher criterion performance may be differently constituted depending on which student learning outcomes are to be obtained.

The multiple regression analysis permitted exploration of the linear relationship between teacher behavior and student learning and this data is reported in table 11. At least one negative correlation suggests a non-linear relationship between teacher behavior and student learning (Table 12). Follow-up study of these relationships supported the contention that many of the student performances appeared to have a non-linear relationship with teacher behavior and that it might be possible to establish teacher criterion behavior as that point at which the relationship between teacher behavior and student learning becomes negative.

For the purposes of this study the data in Tables 10, 11, and 12, support the strong relationship that exists between acquired teacher behavior and student behavior, and the membership of a teacher in a treatment group vs. a control group and student performance.

The foregoing data and discussion indicates that it is possible to combine the validation of teacher behavior research with teacher training research to maximize data collection, knowledge, and cost efficiency. The data reported here provides support for the efficacy of the training methods considered and also confirms the relationship between acquired teacher behavior and student performance, at least with regards to the variables considered in this study.

References

Bandura, Albert. Principles of Behavior Modification. New York: Holt, Rinehart and Winston, 1969.

Bandura, Albert, (ed.). Psychological Modeling: Conflicting Theories. New York: Aldine Atherton, 1971.

Bandura, Albert and McDonald, F. J. The Influence of Social Reinforcement and the Behavior of Models in Shaping Children's Moral Judgments. Journal of Abnormal and Social Psychology. 1963, 67, 274-281.

Bandura, Albert; Ross, D.; and Ross, S. A. Imitation of Film-Mediated Aggressive Models. Journal of Abnormal and Social Psychology. 1963, 66, 3-11.

Bandura, Albert and Walters, R. H. Social Learning and Personality Development. New York: Holt, Rinehart, and Winston, 1963.

Bloom, B. S., (ed.). Taxonomy of Educational Objectives: Handbook I Cognitive Domain. New York: David McKay, 1956.

Combs, Arthur W. Some Basic Concepts for Teacher Education. The Journal of Teacher Education. Vol. XXIII, No. 3, Fall, 1972.

Elam, S. Performance-Based Teacher Education: What is the State of The Art. American Association of Colleges of Teacher Education, December, 1971.

Elam, S. A Resume of Performance-Based Teacher Education. American Association of Colleges of Teacher Education, March, 1972.

Fuller, F. F. Concerns of Teachers: A Developmental Conceptualization. American Educational Research Journal. 1969, 6, 207-226.

Houston, W. R., Housam, R. B., (eds.). Competency-Based Teacher Education. Palo Alto: Science Research Associates, 1972.

Koran, J. J. Jr. The Relative Effects of Classroom Learning and Subsequent Observational Learning on the Acquisition of Questioning Behavior by Pre-Service Elementary Science Teachers. Journal of Research in Science Teaching. 1969, 6, 217-223.

Koran, J. J. Jr. A Comparison of the Effects of Observational Learning and Self-Rating on the Acquisition and Retention of a Questioning Behavior by Elementary Science Teacher Trainees. Science Education. 1970, 54, 385-389.

Koran, J. J. Jr. A Study of the Effects of Written and Film-Mediated Models on the Acquisition of a Science Teaching Skill by Pre-Service Elementary Teachers. Journal of Research in Science Teaching. 1971, 3, 45-50.

Koran, J. J. Jr., Koran, M. L., McDonald, F. J. Effects of Different Sources of Positive and Negative Information on Observational Learning of a Teaching Skill. Journal of Educational Psychology, Vol. 63, No. 5, 405-410, 1972.

Koran, M. L., Snow, R. E., and McDonald, F. J. The Effects of Trainee Aptitudes on Observational Learning. Journal of Educational Psychology. 1971, 62, 219-228.

Masters, J. C., and Branch, M. N. Comparison of the Relative Effectiveness of Instructions, Modeling, and Reinforcement Procedures for Inducing Behavior Change. Journal of Experimental Psychology. 80, 364-368. 1969.

McDonald, F. J. and Allen, D. W. Training Effects of Feedback and Modeling Procedures on Teacher Performance. (USOE Final Report 6-10-078) Stanford University, School of Education, Stanford, California. 1967.

McDonald, F. J., and Koran, M. L. The Effects of Individual Differences on Observational Learning in the Acquisition of a Teaching Skill. USOE Technical Report. Project No. 81073, Stanford University, Stanford, California, 1967.

Rosenshine, Barak. Interpretive Study of Teaching Behaviors Related to Student Achievement. College of Education, Temple University, Philadelphia, Pennsylvania. BR-9B-010, OEG-2-9-480010-1016(010) 163 pages. 1970.

Rosenshine, B., Furst, N. Research in Teacher Performance Criteria. In B. O. Smith, (ed.). Research in Teacher Education: A Symposium. Prentice Hall, 1971.

Rosenshine, Barak, Furst, Norman. The Use of Direct Observation to Study Teaching, in Travers, Robert M. W., (ed.). Second Handbook of Research on Teaching. Chicago: Rand McNally and Company, 1973.

Rosner, Benjamin, Kay, Patricia M. Will the Promise of C/PBTE Be Fulfilled? Phi Delta Kappan, Vol. LV, No. 5, January, 1974.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Table 1
Model for Research*

Phase 1
Training Phase

Treatments	Criterion Measures	Delayed Measures
1. Written Model: Protocol Form N = 24	Microteaching: Audiotape Interaction - Teacher ⇌ Student	Student Written Transfer Measures
2. Written Model: Transcript Form N = 22	Written: Teacher and Student Identification and Discrimination of Analytic Questions	same
3. Control N = 23	Student only: Content Test Affective Test	same

---

Phase 2
Analysis Phase

Treatments	Criterion Measures			
Protocol Transcript Control	Anova Teacher Written Measure Teacher Verbal Behavior Categories Frequency	Anova Student Written Measure Student Verbal Behavior Categories Frequency	Student Criterion	Relationships Regression Analysis
	Correlations			Teacher Criterion

---

*This design incorporates mechanisms for training all teacher exposed to the treatments to criterion on the behavior. Any theoretical and empirically justified treatments may be employed. Maximum data may be generated to answer questions on positive relationships to student learning and on negative effects of teachers at a low level of performance on students.

TABLE 2

DESIGN AND TREATMENT PROCEDURES

Steps	Group			Time
	Protocol Model	Interaction Model	Control	
1. Set Induction	X	X	X	30 minutes
2. Treatment	Protocol	Interaction	Control	
Lesson Content	X	X	X	
3. Microteaching	X	X	X	20 minutes
4. Test Administration	X	X	X	20 minutes
Teacher Tests	X	X	X	
Student Tests	X	X	X	
5. Transfer Test				
Students Only	X	X	X	55 minutes

Explanation of Symbols: (X) indicates that all Ss received this step of the treatment in an identical manner. Written descriptions are provided for the two steps in which treatments varied among the three groups.

Table 3
Reliability of Criterion Measures

Microteaching	Written Measures
[1]Teachers	[3]Teachers
Total Number of Analytic Questions. . . . r = .99	Recognition and Identification Analytic Questions
Number of High Quality Questions. . r = .89	26 Items. . . . . r = .70
Number of Low Quality Questions . . r = .97	[4]Students
Number of Categories of Analytic Questions . . . . . r = .97	Knowledge of Content of Lesson Communication
[2]Students	20 Items. . . . . r = .69
Total Number of Analytic Responses. . . . r = .95	Recognition and Identification of Analytic Questions
Number of Categories of Analytic Responses . . . . . r = .98	32 Items. . . . . r = .71
	Transfer Test-Analysis of New Communication. . . . . r = .79

[1]Number of Teachers = 61
[2]Number of Students = 225-260
[3]Number of Teachers = 67
[4]Number of Students = 269-295

Table 4
Means and Standard Deviations of Dependent Variables

Criterion Measures	Treatment					
	Protocol[a]		Transcripts[b]		Control[c]	
	M	SD	M	SD	M	SD
Teacher Written Measure Identifying Analytic Questions	11.88	4.30	13.77	3.87	9.21	2.94
Student Written Measures Content	13.44	3.32	12.44	4.20	13.21	3.30
Identifying Analytic Questions	9.08	3.37	8.96	3.37	7.50	3.83
Transfer Test	14.99	4.73	14.16	6.63	14.08	5.05
Teacher Tape Measures Number of Analytic Questions	9.14	11.67	4.11	4.92	.95	2.64
High Quality	3.59	4.26	2.58	3.22	.35	1.14
Low Quality	5.50	7.61	1.47	1.84	.55	1.43
Number of Categories	2.00	1.41	1.58	1.43	.30	.80
Student Tape Measures Number of Analytic Responses	5.73	6.97	1.89	2.42	.65	1.69
Number of Categories of Responses	1.77	1.41	1.16	1.21	.25	.72

[a] Tn = 24; Sn = 89, 78
[b] Tn = 22; Sn = 79, 70
[c] Tn = 23; Sn = 85; 73

Table 5
Analysis of Variance for Treatment Groups
Teacher Written Criterion Measure: Distinguishing
Analytic Questions from Non-Analytic

Source	df	Sum Squares	Mean Squares	F
Between Groups	2	236.2360	118.118	8.40**
Within Groups	66	928.4041	14.067	
Total	68	1164.64		

N = 69
**p < .01

Table 6
Analysis of Variance: Student Written Criterion Measure

Source	df	Student Content Test		Student Identification of Analytic Questions	
		MS	F	MS	F
Between Groups	2	22.37	1.71	65.31	5.24**
Within Groups	250	13.06		12.47	
Total	252				

Table 7
Analysis of Variance: Student Written Transfer Test

		Sum of Squares	Mean Squares	F
Between Groups	2	38.25	19.13	0.63
Within Groups	218	6589.76	30.23	
Total	220	6628.01		

**$p < .01$

Table 8
Analysis of Variance: Microteaching Performance Measures
Teachers and Students

Source	df	Teacher Analytic Questions								Student Responses			
		Total		High Qual.		Low Qual.		Categories		Total		Categories	
		MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
Between Groups	2	359.43	6.08**	56.86	5.57**	146.57	6.43**	16.13	10.30**	148.12	7.27**	12.23	9.07**
Within Groups	58	59.16		10.22		22.78		1.57		20.36		1.34	
Total	60												

*p .05
**p .01

Table 9
Summary of Affective Responses

Questions	T_1 a	T_1 b	T_1 c	T_1 d	T_2 a	T_2 b	T_2 c	T_2 d	T_3 a	T_3 b	T_3 c	T_3 d
1. I feel the material contained in this lesson was _____a) very interesting _____b) of some interest _____c) of little or no interest	34	47	7		29	39	1		42	38	3	
2. In order to understand what I read, I think that knowing how to analyze written material into its hypotheses, definitions, assumptions, facts or opinions and conclusions is _____a) absolutely required _____b) extremely important _____c) of some importance _____d) of little importance	11	40	30	7	13	25	29	7	11	29	35	8
3. Think back to the kinds of questions asked by your teacher in this lesson. In understanding the material, these questions were _____a) very helpful _____b) of some help _____c) of little or no help	54	28	6		38	31	5		49	32	2	
4. The questions asked by the teacher in this lesson _____a) forced me to think most of the time _____b) required some thought on my part _____c) enabled me to get the correct answer without thinking hard _____d) required practically no thought on my part	25	45	14	4	14	40	14	7	23	45	11	4
5. I would have understood the material better if _____a) the teacher asked more questions _____b) the teacher asked fewer questions _____c) the lesson was satisfying as it was	27	7	54		34	7	33		19	8	56	
6. In comparing this lesson with the usual lesson, I felt I learned _____a) more efficiently _____b) as efficiently as from a typical lesson _____c) less efficiently than from a typical lesson	50	26	12		32	32	11		39	40	4	

Table 10
Correlations of Independent and Dependent Variables

Variable	1	2	3	4	5	6	7	8	9	10
1. Teacher Written Test: Identification	1.00	0.53	0.48	0.30	0.39	-0.01	0.06	-0.04	0.51	0.33
2. Teacher Categories of Questions		1.00	0.88	0.76	0.84	-0.02	0.21	-0.02	0.94	0.81
3. Teacher Hi Quality			1.00	0.84	0.94	-0.01	0.14	0.05	0.89	0.89
4. Teacher Lo Quality				1.00	0.97	0.07	0.14	0.06	0.80	0.97
5. Teacher Frequency of Questions					1.00	-0.04	0.14	0.06	0.87	0.97
6. Student Content Test						1.00	0.19	0.04	0.45	0.06
7. Student Written Test: Identification							1.00	0.14	0.23	0.16
8. Student Transfer Test								1.00	-0.01	0.07
9. Student Categories of Questions									1.00	0.88
10. Student Total Frequency of Questions										1.00

N = 225
*p < .05 = .14
**p < .01 = .18

Table 11

Multiple Regression Analysis: Independent and Dependent Variables

Independent and Dependent Variables	Regression Coefficients — Student				
	Content Test	Transfer Test	Id. Analytic Questions	Number of Analytic Response Categories	Frequency of Analytic Responses
Teacher Identification of Analytic Questions					-.06
Teacher Number of Analytic Question Categories			.51	.64	.31
Teacher H Q Questions					
Teacher L Q Questions					.40
Teacher Total Questions				.04	.31
R			.21	.95	.98
R^2			.04	.90	.96
Intercept			7.70	.02	.58
Signif. of Prediction (F)			9.85** (2,223)	1005.85** (2,222)	1249.14** (4,220)

*p < .05 **p < .01